

DEW



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Foster D. Hinshaw, Vincent F. Femia, Craig S. Harris, John K. Metzger,
David L. Meyers, and Barry M. Zane

Application No.: 10/667,127 Group Art Unit: 2651

Filed: September 18, 2003 Examiner: Not assigned

Confirmation No.: 7163

Title: DISK MIRROR ARCHITECTURE FOR DATABASE APPLIANCE

CERTIFICATE OF MAILING OR TRANSMISSION	
I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as First Class Mail in an envelope addressed to Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, or is being facsimile transmitted to the United States Patent and Trademark Office on:	
<u>11/10/04</u>	<u>Ellen T. Spear</u>
Date	Signature
<u>Ellen T. Spear</u>	
Typed or printed name of person signing certificate	

11/15/2004 LWONDIM1 00000015 10667187
01 FC:1460

PETITION TO MAKE SPECIAL FOR NEW APPLICATION
UNDER M.P.E.P. § 708.02, VII

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Applicants hereby petition to make this new application, which has not received any examination by an Examiner, special as permitted under 37 CFR § 1.102(d). The Petition fee in the amount of \$130 is included in the enclosed check. Please charge any deficiency in fees and credit any overpayment to Deposit Account 08-0380. A duplicate copy of this Petition is enclosed for accounting purposes.

Applicants believe that all the claims in this application are directed to a single invention. If the Office determines that all the claims presented are not obviously directed to a single invention, then applicants will make an election without traverse as a prerequisite to the grant of special status.

A pre-examination search has been made by the inventors and their attorneys. The search strategy included reviewing publications known to the inventors and additional searching using terms such as “database”, “disk mirroring”, “processing unit”, “data redundancy”, “data mirroring” and “disk redundancy”.

A Supplemental Information Disclosure Statement and accompanying PTO-1449 form are being filed concurrently with this Petition To Make Special.

Below is a detailed discussion of the present invention and the prior art references, which particularly points out how the claimed subject matter is distinguishable.

The present invention is a disk mirror system that includes a plurality of processing assemblies, each consisting of one or more disks and a processing unit. Each disk has at least two data segments, a first data segment and one or more secondary data segments, and may have one or more system segments. Each processing unit is coupled to one or more of the plurality of disks.

A processing unit that receives a request to write data to a first disk writes the data to the first data segment of the first disk and forwards the data to another processing unit. The data set on the first disk is referred to as the primary data segment or the primary data slice. The other processing unit writes a copy of the data to a secondary data segment of a second disk coupled to

the other processing unit. The data set copied to the second disk is referred to as the mirror data set or the mirror data slice.

The first data segment may include fast tracks of the disk and the secondary data segment may include slow tracks of the disk. The fast tracks may be the outer tracks of the disk and the slow tracks may be the inner tracks of the disk. The disks need not mutually mirror each other. That is, there may be multiple disks and a first segment of a first disk may be mirrored on a second disk, while a first segment of a second disk may be mirrored on a third disk.

U.S. Patent No. 6,801,921 to Tsuchida, *et al.* discloses a database system with multiple storage units. The database is duplicated for redundancy and the system enables maintenance and reorganization of database storage units without interrupting the ongoing operations involving access to the database storage units. The non-stop ongoing database access is provided by allocating a pair of a main storage unit and an intermediate storage unit that maintains the substantially identical data. During certain predetermined operations, the main storage unit is dedicated to the ongoing transactions while the intermediate database storage unit is manipulated to contain data that corresponds to a specified time. For example, the content of the intermediate storage is subsequently copied or used to update another database. When the above duplication or incorporation is completed, the intermediate storage unit is updated for the transactions that have taken place during the above operations to be synchronized with the main storage unit.

While Tsuchida, *et al.* does disclose multiple database storage units and duplicating some data across two storage units, it does not teach multiple processing units or disks associated with processing units. Nor does it teach logically dividing storage of the disk into at least two segments and using one of the data segments for a primary storage and the second data segment as a mirror of the primary data segment of another disk.

U.S. Patent No. 6,405,284 to Bridge describes a system for managing storage from multiple data storage devices. Each disk drive is divided into many small fixed size pieces, which are mirrored on other disks. As disk configuration and availability changes, data from some fixed size pieces may be moved or copied onto other disk drives. In another patent to Bridge, U.S. Patent No. 6,530,035, disks are associated with failure groups. Two disk drives are in different failure groups if they do not share a common failure condition that is projected to affect both disk drives at the same time. For mirroring, each disk drive is paired with one or more disk drives from other failure groups. These symmetric pairs are called "mirror partners." Pairs are formed so that every disk drive has mirror partners in as many different failure groups as possible. When a piece of a logical volume is allocated on a disk drive, the mirror copy of that piece is allocated on one of the disk drive's mirror partners. The partner is chosen to evenly distribute the mirrored data to all partners.

The system of Bridge does contain multiple disks, with sections of the disk being mirrored on other disks. However, Bridge does not disclose multiple processing units, each processing unit being associated with a disk.

U.S. Patent No. 5,764,903 to Yu discloses a system for mirroring hard disk data of a primary server over a network to a secondary server. A virtual disk driver is provided between the operating system and the hard disk driver of the primary server. After the virtual disk driver receives a disk write request that should be mirrored, the virtual disk driver first sends a disk write request to the secondary server. The virtual disk driver then sends the disk write request to a disk drive on the primary server. In the event data residing on one server is unavailable due to a system failure, the same data would still be available on the other server.

Yu does not teach logically dividing storage of one disk – what would be a primary disk in the system of Yu – into at least two segments and using one of the data segments for a primary

storage and the second data segment as a mirror of the primary data segment of another disk. Instead, Yu merely mirrors one primary disk on one or more secondary disks.

U.S. Patent No. 6,694,406 to Kodama, *et al.* relates to a method of allocating each of a number of processor units to a corresponding one of a number of disk storage units. In this way, each processor unit can read data from its allocated disk storage unit with minimum conflict to other read and/or write operations conducted at or about the same time by other processor units. The processing system of Kodama *et al.* includes a number of processor elements connected to disk storage having a plurality of disk storage units for maintaining data. One of the processor elements, designated a "Mount Manager," is responsible for assigning a disk storage unit to a corresponding one of the other processor elements so that, preferably, there is a one-to-one correspondence between a disk storage unit and a processor element. One of the disk storage units is designated a master disk unit, and the remaining disk storage units are designated "mirrored" disk units.

The processing units of Kodama, *et al.* are processors requesting read and/or write access to data. Write access is directed to the main disk, while read access commands are directed to associated mirroring disks. Kodama, *et al.* does not teach disks subdivided into two or more segments, different segments on different disks mirroring each other, nor is association between the storage units and processing units like in the present application, where the processing unit controls access to the storage disk.

U.S. Patent No. 4,077,059 to Cordi, *et al.* discloses a multi-processing system with a hierarchical memory having journaling and copyback. The hierarchical memory system has two memory units on each level. One main unit contains all of the data for the level, and the other unit, the copyback unit, contains the changes that have been made to that data either by addition or modification. The main unit interfaces with the next higher level in the hierarchy and its

processing unit. The copyback unit transfers the data changes to the lower level when the lower level's main unit is not interfacing with its next higher level or processing unit. The copyback unit is smaller than the main unit to reduce the necessary storage units on each level.

The two-stage memory system of Cordi, *et al.* is unlike the mirroring system of the present application, because it does not create different segments mirroring each other. Instead, it deals with storing updates to one memory unit in another memory unit.

U.S. Patent No. 6,389,459 B1 to McDowell discloses a system for mirroring data and disk drive device state over a network between a source computer including a first non-shared disk drive storage device and a target computer including a second non-shared disk drive storage device. The system includes layered software drivers installed within the source computer above the device driver for the first disk drive storage device for capturing and examining write requests directed to the first disk drive storage device, and transmitting the write request and extracted device state information to the target computer.

The mirroring system of McDowell is unlike the mirroring system of the present application, because it does not mirror segments of one disk on another disk, while mirroring some other disk on the first disk. Instead, it is a standard mirroring system where one disk is completely mirrored on another disk, and the process of copying data takes place over the network.

U.S. Patent No. 6,298,425 to Whitaker, *et al.*, describes a system that stores two copies of the metadata in two areas on disk. Each individual block of metadata in each area is duplicated as a doublet. This enables metadata to survive a single block failure since such a failure only affects half of the doublet. Further, each block of metadata has header information including an

indication of transaction and a part count. A transaction is not deemed committed until corresponding blocks of all parts are found.

The system of Whitaker, *et al.* essentially does data mirroring, but only on a single disk. There is no suggestion of replicating data across multiple disks associated with multiple processing units.

U.S. Patent No. 6,606,694 B2 to Carteau deals with a system in which disk drives are mirrored through duplication controlled by disk controllers. Each disk controller controls writing to a set of disk drives. A disk write request to one disk controller causes that disk controller to write to one of its disks and to transmit the write request to another controller that in turn writes to its disk. The second controller then acknowledges the write to the first controller, which in turn acknowledges the write to the computer issuing the request. The first controller further logs the writes in a log file. This logging allows efficient resynchronization after mirroring is broken and reestablished, as well as removing cable length restrictions between controllers.

Carteau describes a standard mirroring setup, but with a logging feature. This mirroring system is unlike that of the present application because it does not create different disk segments on one disk, one segment being a primary data storage, while another mirroring a primary segment from another disk.

U.S. Patent No. 6,654,862 to Morris discloses a set of storage disks that includes a mirrored pair of disks. Each disk in the pair includes a mirror copy of data stored on the other disk in the pair. For each of the disks in the pair, the data forming the mirror copy is physically located between the center of the disk and the data that is mirrored on the other disk. Morris does not teach disks being coupled to processing units, as in the present application.

U.S. Application No. 20020144068 to Ohran discloses a system for updating a mirror of a primary storage device through tracking changes made to the primary mass storage device. These changes are consolidated periodically into update files, the consolidations representing changes made to the primary mass storage during a time interval that ends when the primary mass storage is in a logically consistent state. These update files contain only those changes necessary to represent the modified state of the primary mass storage at the time of the update. The primary system then transfers the update files to a secondary system to bring the secondary mass storage current with the primary mass storage. Ohran does not teach disks subdivided into two or more segments, different segments on different disks mirroring each other.

U.S. Patent No. 6,792,486 to Hanan, *et al.* describes a system for managing information storage among plural disk drives. The system includes plural host interfaces and first and second elements. The first and second elements each comprise a set of disk drives for storing information. Each of the first and second elements is associated with an element frame. Each disk drive included in the first element is connected to a different one of the plural host interfaces during the element frame of the first element. The system includes a switch controller for selecting among the first and second elements, and for directing information from a first one of the plural host interfaces to a selected disk drive within the first element.

Hanan, *et al.* does not teach a set of disks partitioned into two or more segments, the disks mirroring each other. The host interfaces of Hanan, *et al.* are also different from the processing units of the present application.

U.S. Patent No. 5,828,820 to Onishi, *et al.*, describes a system in which specific numbers identical to those written in disks and name of a drive used as for a master disk are recorded in a set-up memory incorporated therein, and this data is compared to a specific number written in a disk to determine a master disk or to transfer data to a slave disk. In such a way, a master disk is

clearly differentiated from a slave disk, and, in case of a disk failure, a recovery process can proceed automatically. Onishi, *et al.* does not teach disks subdivided into two or more segments, nor does it disclose different processing units being associated with each disk.

U.S. Patent No. 5,737,601 to Jain, *et al.*, discloses methods for replicating modifications made at a local site to multiple remote sites in a peer-to-peer environment. Information regarding these replicated modifications (e.g., insert, delete, or update) is contained in a set of replication tables. Modifications can be duplicated at other sites. Jain, *et al.* further provides a logic-oriented procedure-level replication. Procedure-level replication modifies a remote site based on the logical operations used to modify the data at the originating site.

Jain, *et al.* does not deal with physical organization of the storage units, nor does it disclose a mirroring system where one segment of one disk mirrors a second segment of a second disk. The replication methods of Jain, *et al.* are implemented preferably in software and not firmware, as in the present application.

The following references generally deal with keeping multiple versions of a database updated with respect to each other. None of them disclose the physical disks being divided into at least two segments, one segment storing primary data, and the second segment storing mirror data of another segment of another disk.

U.S. Patent No. 4,648,036 to Gallant discusses a method for controlling query and update processing in a database system. Specifically, Gallant describes a method for ensuring that a query receives information representative of the database either before or after an update, but not information representative of a state after the update begins but before it completes. Transactional modifications are made to a future database structure. At the completion of the

modifications, a switch is made from the present database structure to the future database structure. A query process accesses the present database structure.

U.S. Patent No. 4,646,229 to Boyle describes a database system that includes future versions of the database for use in time-oriented applications such as an application for scheduling the use of the same facilities to present and future users. All of the information required to represent the data base contents at desired future points in time is maintained in the data base.

U.S. Patent No. 4,714,992 to Gladney, *et al.*, discloses a method for managing obsolescence of replicas of data objects in a distributed processing system. Database objects at a source location are replicated at a replica location. As objects stored at the source location are altered, corresponding objects at the replica location become obsolete. A replica location generates a request for a list of obsolete objects from the source location.

U.S. Patent No. 4,635,189 to Kendall discloses a real-time distributed database management system that stores in a processor's memory copies of the variables needed to run the programs in that processor. When a variable is created, a processor is designated as the processor that determines the value of that variable. Each variable copy is updated by the current value of the original value on a periodic basis or upon the occurrence of a defined condition.

U.S. Patent No. 4,853,843 to Ecklund, *et al.*, describes a multi-version database where each update operation creates a new version of the database, and the older versions remain available. Multiple alternative version paths are retained.

The following references are only of more general interest.

U.S. Patent No. 6,167,531 to Sliwinski deals with methods for transferring sets of mirrored disks from an active system to a standby system in the event of a failure of the active system.

U.S. Patent No. 6,801,914 issued to Barga, *et al.* describes a client-server database system that enables persistent client-server database sessions, without modification to a client-side application, a database system, or native client-side database drivers.

U.S. Patent No. 4,631,673 to Haas, *et al.*, discloses a method for refreshing multicolumn tables in a relational database. Haas provides a method for refreshing the a read-only copy of a base table portion.

U.S. Patent No. 4,558,413 to Schmidt, *et al.*, discloses a management system for managing versions of software updated and stored in designated storage devices in a distributed software environment of a local area network.

U.S. Patent No. 6,023,584 to Barton, *et al.*, discloses methods for installing a computer program from a removable data storage device using disk-mirroring techniques.

U.S. Patent No. 4,875,159 to Cary, *et al.*, describes a system for synchronizing two versions of files in a multiprocessor system.

These references are believed only to be of general interest since they do not disclose a system that includes a processing unit coupled to one or more disks, and where each disk has two or more segments, one segment being the primary data segment, and another segment mirroring a primary segment of another disk.

In summary, none of the above patents disclose the system, apparatus, method or computer program product of the present invention. We therefore ask that this Petition to Make Special for the present application be granted.

Respectfully submitted,

HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

By *Lubashev*
Lyudmila Lubashev
Registration No. 55,408
Telephone: (978) 341-0036
Facsimile: (978) 341-0136

Concord, MA 01742-9133
Dated: *11/10/04*

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Foster D. Hinshaw, Vincent F. Femia, Craig S. Harris, John K. Metzger, David L. Meyers, and Barry M. Zane

Application No.: 10/667,127

Group Art Unit: 2651

Filed: September 18, 2003

Examiner: Not assigned

Confirmation No.: 7163

Title: DISK MIRROR ARCHITECTURE FOR DATABASE APPLIANCE

CERTIFICATE OF MAILING OR TRANSMISSION	
I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as First Class Mail in an envelope addressed to Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, or is being facsimile transmitted to the United States Patent and Trademark Office on:	
11/10/04	<i>Ellen T. Spear</i>
Date	Signature
<i>Ellen T. Spear</i>	
Typed or printed name of person signing certificate	

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

This Information Disclosure Statement is submitted:

☐ under 37 CFR 1.129(a), or
(First/Second submission after Final Rejection)

☒ under 37 CFR 1.97(b), or
(Within any one of the following time periods: three months of filing national application (other than a CPA) or date of entry of the national stage in an international application; or before the mailing date of a first office action on the merits in a non-provisional application, including a CPA, or a Request for Continued Examination).

☐ under 37 CFR 1.97(c) together with either:
☐ a Statement under 37 CFR 1.97(e), as checked below, or
☐ a \$180.00 fee under 37 CFR 1.17(p), or
(After the 37 CFR 1.97(b) time period, but before final action or notice of allowance, whichever occurs first)

☐ under 37 CFR 1.97(d) together with:
☐ a Statement under 37 CFR 1.97(e), as checked below, and
☐ a \$180.00 fee under 37 CFR 1.17(p), or
(Filed after final action or notice of allowance, whichever occurs first, but on or before payment of the issue fee)

☐ under 37 CFR 1.97(i):
Applicant requests that the IDS and cited reference(s) be placed in the application filewrapper.
(Filed after payment of issue fee)

Statement Under 37 CFR 1.97(e)

- ☐ Each item of information contained in this Information Disclosure Statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of this Information Disclosure Statement; or
- ☐ No item of information contained in this Information Disclosure Statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the undersigned, after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of this Information Disclosure Statement.

Statement Under 37 CFR 1.704(d) (Patent Term Adjustment)

Applies to original applications (other than design) filed on or after May 29, 2000

- ☐ Each item of information contained in the Information Disclosure Statement was cited in a communication from a foreign patent office in a counterpart application and this communication was not received by any individual designated in § 1.56(c) more than thirty days prior to the filing of the Information Disclosure Statement.
- ☒ Enclosed herewith is form PTO-1449:
- ☐ Copies of the cited references are enclosed.
- ☒ Since this application was filed after June 30, 2003, copies of issued U.S. patents and published U.S. applications are not required and are not being provided.
- ☐ Copies of the cited references are enclosed except those entered in prior application, U.S. Application No. [], to which priority under 35 U.S.C. 120 is claimed. [The earlier application contains copies of the cited references.]
- ☐ The listed references were cited in the enclosed International Search Report in a counterpart foreign application.
- ☐ The "concise explanation" requirement (non-English references) for reference(s) [] under 37 CFR 1.98(a)(3) is satisfied by:
- ☐ the explanation provided on the attached sheet.
- ☐ the explanation provided in the Specification.
- ☐ submission of the enclosed International Search Report.
- ☐ submission of the enclosed English-language version of a foreign Search Report and/or foreign Office Action.
- ☐ the enclosed English language abstract.

☐ Applicant requests that the following non-published pending applications be considered:

Examiner's
Initials

____ U.S. Patent Application No. [], by [inventor(s)], filed [], Docket No.: []
____ U.S. Patent Application No. [], by [inventor(s)], filed [], Docket No.: []
____ U.S. Patent Application No. [], by [inventor(s)], filed [], Docket No.: []

Examiner

Date

- ☐ A copy of each above-cited application, including the current claims, is enclosed.
- ☐ A copy of each above-cited application, including the current claims, is enclosed, except those entered in prior application, U.S. Application No. [], to which priority under 35 U.S.C. 120 is claimed.

The Examiner is requested to return a copy of the above list of pending applications indicating which references were considered with the next office communication.

It is requested that the information disclosed herein be made of record in this application.

Method of payment:

- ☐ A check for the fee noted above is enclosed, or the fee has been included in the check with the accompanying Reply. A copy of this Statement is enclosed.
- ☐ Please charge Deposit Account 08-0380 in the amount of \$[]. A copy of this Statement is enclosed.
- ☒ Please charge any deficiency in fees and credit any overpayment to Deposit Account 08-0380.

Respectfully submitted,

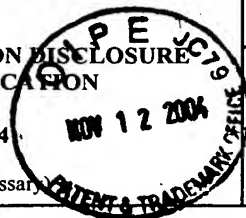
HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

By Lubashev
Lyudmila Lubashev
Registration No.: 55,408
Telephone: (978) 341-0036
Facsimile: (978) 341-0136

Concord, MA 01742-9133

Dated: 11/10/04

PTO-1449 REPRODUCED		ATTORNEY DOCKET NO. 3336.1001-001		APPLICATION NO. 10/667,127	
SUPPLEMENTAL INFORMATION DISCLOSURE CITATION IN AN APPLICATION November 10, 2004 (Use several sheets if necessary)		FIRST NAMED INVENTOR Foster D. Hinshaw		FILING DATE September 18, 2003	
		EXAMINER Not Assigned		CONFIRMATION NO. 7163	
				GROUP 2651	



U.S. PATENT DOCUMENTS				
EXAM- INER INI- TIAL	REF. NO.	DOCUMENT NUMBER Number-Kind Code (if known)	ISSUE DATE / PUBLICATION DATE MM-DD-YYYY	NAME OF PATENTEE OR APPLICANT OF CITED DOCUMENT
	AG	US 6,801,921 B2	10/05/2004	Tsuchida et al.
	AH	US 6,405,284 B1	06/11/2002	Bridge
	AI	US 6,530,035 B1	03/04/2003	Bridge
	AJ	US 6,694,406 B2	02/17/2004	Kodama et al.
	AK	4,077,059	02/28/1978	Cordi et al.
	AA2	US 6,298,425 B1	10/02/2001	Whitaker et al.
	AB2	US 6,654,862 B2	11/25/2003	Morris
	AC2	US 6,792,486 B1	09/14/2004	Hanan et al.
	AD2	5,737,601	04/07/1998	Jain et al.
	AE2	4,648,036	03/03/1987	Gallant
	AF2	4,646,229	02/24/1987	Boyle
	AG2	4,714,992	12/22/1987	Gladney et al.
	AH2	4,635,189	01/06/1987	Kendall
	AI2	4,853,843	08/01/1989	Ecklund et al.
	AJ2	US 6,801,914 B2	10/05/2004	Barga et al.
	AK2	4,631,673	12/23/1986	Haas et al.
	AA3	4,558,413	12/10/1985	Schmidt et al.
	AB3	4,875,159	10/17/1989	Cary et al.
	AC3	US 2002/0144068 A1	10/03/2002	Ohran

EXAMINER	DATE CONSIDERED
----------	-----------------